

AD\_\_\_\_\_

Award Number: Contract W81XWH-12-2-0070

TITLE: Annual Report: The Assessment of Military Multitasking Performance: Validation of a Dual-Task and Multitask Protocol (Contract W81XWH-12-2-0070)

PRINCIPAL INVESTIGATOR: Margaret M. Weightman PT, Ph.D.

CONTRACTING ORGANIZATION: Allina Health System  
Minneapolis, MN 55407-3723

REPORT DATE: August 2014

TYPE OF REPORT: Annual Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation

| Report Documentation Page  |                                    |                                     | Form Approved<br>OMB No. 0704-0188                               |   |                                    |
|--|------------------------------------|-------------------------------------|--|---|------------------------------------|
| Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. |                                    |                                     |  |   |                                    |
| 1. REPORT DATE<br><b>AUG 2014</b>  |                                    | 2. REPORT TYPE                      |  | 3. DATES COVERED<br><b>15-08-2013 to 14-08-2014</b> |                                    |
| 4. TITLE AND SUBTITLE<br><b>The Assessment of Military Multitasking Performance: Validation of a Dual-Task and Multitask Protocol</b>  |                                    |                                     | 5a. CONTRACT NUMBER  |   |                                    |
|  |                                    |                                     | 5b. GRANT NUMBER   |   |                                    |
|  |                                    |                                     | 5c. PROGRAM ELEMENT NUMBER                                       |   |                                    |
| 6. AUTHOR(S)   |                                    |                                     | 5d. PROJECT NUMBER   |   |                                    |
|  |                                    |                                     | 5e. TASK NUMBER  |   |                                    |
|  |                                    |                                     | 5f. WORK UNIT NUMBER   |   |                                    |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br><b>Alina Health System,,Minneapolis,MN,55407</b>   |                                    |                                     | 8. PERFORMING ORGANIZATION<br>REPORT NUMBER                      |   |                                    |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)  |                                    |                                     | 10. SPONSOR/MONITOR'S ACRONYM(S)                                 |   |                                    |
|  |                                    |                                     | 11. SPONSOR/MONITOR'S REPORT<br>NUMBER(S)                        |   |                                    |
| 12. DISTRIBUTION/AVAILABILITY STATEMENT<br><b>Approved for public release; distribution unlimited</b>  |                                    |                                     |  |   |                                    |
| 13. SUPPLEMENTARY NOTES  |                                    |                                     |  |   |                                    |
| 14. ABSTRACT   |                                    |                                     |  |   |                                    |
| 15. SUBJECT TERMS  |                                    |                                     |  |   |                                    |
| 16. SECURITY CLASSIFICATION OF:  |                                    |                                     | 17. LIMITATION OF<br>ABSTRACT<br><b>Same as<br/>Report (SAR)</b> | 18. NUMBER<br>OF PAGES<br><b>24</b>                 | 19a. NAME OF<br>RESPONSIBLE PERSON |
| a. REPORT<br><b>unclassified</b>   | b. ABSTRACT<br><b>unclassified</b> | c. THIS PAGE<br><b>unclassified</b> |  |   |                                    |

| REPORT DOCUMENTATION PAGE  |                  |                                 |                            | Form Approved<br>OMB No. 0704-0188                  |   |
|--|------------------|---------------------------------|----------------------------|---|---|
| Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. <b>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</b>  |                  |                                 |                            |   |   |
| 1. REPORT DATE<br>August 2014  |                  | 2. REPORT TYPE<br>Annual Report |                            | 3. DATES COVERED<br>15 August 2013 – 14 August 2014 |   |
| 4. TITLE AND SUBTITLE<br><br>The Assessment of Military Multitasking Performance: Validation of a Dual Task and Multitask Protocol   |                  |                                 |                            | 5a. CONTRACT NUMBER                                 |   |
|  |                  |                                 |                            | 5b. GRANT NUMBER<br>W81XWH-12-2-0070                |   |
|  |                  |                                 |                            | 5c. PROGRAM ELEMENT NUMBER                          |   |
| 6. AUTHOR(S)<br>Margaret M. Weightman, PT, PhD<br><br>E-Mail: margaret.weightman@allina.com  |                  |                                 |                            | 5d. PROJECT NUMBER                                  |   |
|  |                  |                                 |                            | 5e. TASK NUMBER                                     |   |
|  |                  |                                 |                            | 5f. WORK UNIT NUMBER                                |   |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br><br>Allina Health System<br><br>Minneapolis, MN 55407-3723   |                  |                                 |                            | 8. PERFORMING ORGANIZATION REPORT NUMBER            |   |
| 9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)<br>U.S. Army Medical Research and Materiel Command<br>Fort Detrick, Maryland 21702-5012  |                  |                                 |                            | 10. SPONSOR/MONITOR'S ACRONYM(S)                    |   |
|  |                  |                                 |                            | 11. SPONSOR/MONITOR'S REPORT NUMBER(S)              |   |
| 12. DISTRIBUTION / AVAILABILITY STATEMENT<br>Approved for Public Release; Distribution Unlimited   |                  |                                 |                            |   |   |
| 13. SUPPLEMENTARY NOTES  |                  |                                 |                            |   |   |
| 14. ABSTRACT<br>The primary objective of this project is to further refine and conduct preliminary validation of a novel set of test-tasks known as the <i>Assessment of Military Multitasking Performance</i> (AMMP). The AMMP is a battery of functional dual-tasks and multitasks that simulate the combined sensorimotor, cognitive, and exertional demands of Soldiering. Ultimately, it may be used in combination with other metrics to inform duty readiness decisions after concussion/mild traumatic brain injury (mTBI). Regulatory approvals were obtained and data collection began at Fort Bragg in August 2013 with a total of 60 subjects tested to date (36 with mTBI and 24 healthy control). Using an iterative process based on data from both healthy controls and subjects with mTBI, task refinement has resulted in test tasks with overall good to excellent interrater reliability. As of January 2014, data collection has focused on determining preliminary construct and discriminate (known groups) validity of the AMMP. Results of this study will establish whether or not the AMMP can distinguish between Soldiers with mTBI symptom complex and healthy controls; further investigation is needed to determine its ability to predict readiness for return to duty. |                  |                                 |                            |   |   |
| 15. SUBJECT TERMS<br>concussion, mild traumatic brain injury, dual-task assessment, multitask assessment, return-to-duty decisions   |                  |                                 |                            |   |   |
| 16. SECURITY CLASSIFICATION OF:  |                  |                                 | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES                                 | 19a. NAME OF RESPONSIBLE PERSON           |
| a. REPORT<br>U   | b. ABSTRACT<br>U | c. THIS PAGE<br>U               |                            |   | USAMRMC                                   |
|  |                  |                                 | UU                         | 24  | 19b. TELEPHONE NUMBER (include area code) |

## Table of Contents

### Page

|                                   |    |
|-----------------------------------|----|
| Introduction.....                 | 4  |
| Body.....                         | 4  |
| Key Research Accomplishments..... | 13 |
| Reportable Outcomes.....          | 14 |
| Conclusion.....                   | 15 |
| References.....                   | 15 |
| Appendices.....                   | 16 |

**Annual Report**  
**The Assessment of Military Multitasking Performance: Validation of a Dual-Task and Multitask Protocol (Contract W81XWH-12-2-0070)**

**INTRODUCTION:** The primary objective of this project is to further refine and conduct preliminary validation of a novel set of test-tasks known as the *Assessment of Military Multitasking Performance (AMMP)*. The AMMP is a battery of functional dual-tasks and multitasks that simulate the combined sensorimotor, cognitive, and exertional demands of Soldiering for use after concussion/mild traumatic brain injury (mTBI) (Radomski et al., 2013). Investigators anticipate that future studies will be required to evaluate the responsiveness and predictive validity of the AMMP task battery. It is projected that a fully validated AMMP will discriminate between “duty-ready” and “non-duty ready” military Service Members (SM) following mTBI when used in combination with other metrics to inform duty-readiness decisions. It is further anticipated that a validated AMMP will contribute much needed objectivity to the current return to duty (RTD) determination process (Scherer et al., 2013). We expect that future studies will inform how the AMMP can be leveraged in combination with other psychological, physical, demographic and soldiering metrics to develop a clinical prediction rule for recommendations on RTD in SM with mTBI exposure. This report covers the period from 15August 2013 to 14August 2014 and is an annual report. Focus initially has been on test task refinement (See Table 1) and evaluating inter-rater reliability. We received a no cost extension through 14August 2015 to extend our data collection and analysis following delays in IRB approvals and slower than anticipated subject recruitment.

**Table 1: AMMP Test Task Descriptions and Example Revisions**

| AMMP Task                    | Task Description   | Example Revisions based on IRR findings and rater comments  |
|------------------------------|--|---|
| Charge of Quarters (CQ) Duty | Requires the subject to organize and implement a plan in order to complete a number of tasks all while pulling CQ duty. Tasks such as assembling a footstool, inventorying supplies, radioing accountability information to superiors all occur under time and efficiency rules.   | -Improve operational definitions of “rule breaks” and rules for counting number of times a specific rule break occurs.  |
| PATROL-Exertion              | Subject is challenged to gather information from video surveillance and radio communications while exercising at 65 to 85% of the subject’s age predicted maximal heart rate by doing continuous step-ups on an exercise step to simulate a dismounted patrol. IED markers and pertinent logistical information must be recalled and reported at specific times while also requiring a reaction time trigger switch press to a intermittently occurring tone sound | -Video revisions including change from a SALUTE format to “general Patrol reporting”.<br>-Improved score sheets by grouping responses and clarifying scoring instructions.<br>-Addition of a reaction time task using a trigger switch press in response to a randomly occurring auditory tone. |
| Run-Roll-Aim (RRA)           | Subject completes several typical military maneuvers such as jumping over a trip wire, a 3-5 second rush, combat rolls, side shuttling and back pedaling all while carrying a simulated weapon, the subject uses a short focal point scope on the  | -Improve operational definitions of errors, and ‘hesitations’ during Stroop response.   |

|   |  |  |
|---|--|--|
|   | weapon to identify numbers on a computer screen based on instructions given before the task starts.  |  |
| Load Magazine-Radio Chatter                         | Subject loads M-16 dummy rounds from a bin of mixed size rounds as fast as possible. The dual-task condition requires monitoring radio communication about an upcoming training event and verbally announcing when key words are spoken by specific personnel          | -No modifications to task materials.<br>Require rater to sit close to or in front of subject to ensure clear hearing of call outs for key word indicators.               |
| Illinois Agility-Word List (IAT)                    | The IAT is used in sport assessments; it requires running a course with rapid direction changes and figure of eight navigation of cones. In the dual task condition, the subject is given a 7 word packing list to remember.   | -Modify word-list to remove most compound words or phrases.<br>-Clarify scoring rules for “approximated” but not exact word recalls.                                     |
| Instrumented Stand and Walk-Grid Coordinates (ISAW) | Using wireless wearable inertial sensors and a clinical software program to measure static postural sway and then dynamic stability during walking and turning. A grid memorization task provided in the context of a patrol mission provides the cognitive challenge. | -Clarify scoring rules for recalled grid coordinates.<br>-Require rater to stand in front of subject to ensure clear hearing of call outs for recalled grid coordinates. |

#### **BODY:**

**Aims of the proposal as described in the SOW are:**

**Aim 1: Further specify and refine a set of dual and multitasks with procedures for test administration.**

**Approach:** Task refinement, preliminary retest reliability and scoring testing will be piloted at Courage Kenny Research Center (CKRC) in healthy control SM. Reliability testing at US Army Research Institute of Environmental Medicine (USARIEM) (and/or US Army Aeromedical Research Laboratory (USAARL)) among healthy control SM will further delineate tasks that meet evaluation criteria (**see Table1 on Task Evaluation Criteria**). Additional interrater reliability testing has been added at Fort Bragg/Womack Army Medical Center following additional task revisions.

Tasks that do not meet feasibility or reliability specifications will be dropped or revised. Test-task evaluation will continue throughout the study to examine their ability to discriminate SM with mTBI symptom complex and healthy control SM. Correlations with neurobehavioral testing of known mTBI vulnerabilities (components of ANAM, dynamic visual acuity, tests of selective attention, processing speed and working memory) will be confirmed. Final determination of tasks that remain in the AMMP and the means to combine scoring into a single or multiple metrics will be determined in consultation with test development experts throughout the project.

**Progress toward Aim 1:** Task evaluation is an ongoing process using the criteria in the Task Evaluation Criteria Table (Table 2) which up to this point have included factors such as test burden including test time, scoring objectivity and inter-rater reliability findings. In addition to work described in the first annual report (September 2013), inter-rater reliability data and feedback from the AMMP data collection team have been used to refine scoring metrics, administration scripts, and subject materials.

**Table 2: Task Evaluation Criteria**

| Parameter                           | Description  |
|-------------------------------------|--|
| 1. Total time, set up, take down    | Time   |
| 2. Storage space                    | Square feet  |
| 3. Objective scoring                | Scoring can be constructed to be objective   |
| 4. Dimension(s) challenged          | Which of the dimensions are challenged by the task   |
| 5. Time to give instructions        | Time from beginning instructions to participant beginning the task   |
| 6. Participant's evaluation of task | Questionnaire to obtain feedback about performance of the task(s)  |
| 7. Inter-rater reliability          | Minimum of 0.85 on observations of complex task assessments.<br>Minimum of 0.90 for dual-task assessments.                                     |
| 8. Convergent/discriminant validity | Correlations between task scores and hypothesized related tests of common domain scores will be significant (non-zero) and not less than 0.40. |
| 9. Known groups construct validity  | Significant differences between healthy controls and SM with mTBI complex symptoms at p-value <0.05 and minimum effect size of 0.5.            |

**Implications:** As indicated above, IRR data and ongoing input from data collectors has informed relatively minor but important improvements in scoring and test administration clarity for at least 3 of the 6 test tasks.

These data also informed substantive modifications in what was previously called the SALUTE-Exertion test task. Rank, military occupational specialty and combat experience appeared to unduly influence SM ability to complete a SALUTE report according to AMMP test specifications, creating artifact in terms of potential of this test task to distinguish between groups of SM with mTBI versus healthy controls (HC). Based on feedback from a recently deployed Army subject matter expert, the following test task modifications were made: video revisions including change from a SALUTE format to “general Patrol reporting”; improved score sheets by grouping responses and clarifying scoring instructions. Other changes to this task included the addition of a reaction time task using a trigger switch press in response to a randomly occurring auditory tone. This task was then renamed **PATROL-Exertion** Task for future testing.

**Aim 2: Evaluate inter-rater for each of the dual-tasks and multitasks using healthy control and SM with diagnosed mTBI.**

**Approach:** Inter-rater reliability and assessment of training requirements for expert and novice raters will be completed using 20-25 Human Research Volunteers and permanent party personnel at USARIEM. We anticipate that several test-tasks that do not meet evaluation criteria (see **Table 1**) will be eliminated from the protocol based on initial testing. This will decrease testing time and burden in subsequent data collection. At Fort Bragg/Womack Army Medical Center (WAMC) 80 healthy control and 80 SM with mTBI symptom complex will be tested using the scaled down AMMP. We will assess order effects based on observation of performance as well as actual task scores in the Fort Bragg/WAMC testing cohorts. Inter-rater reliability of the AMMP will be evaluated in a subset of SM with mTBI at Fort Bragg. Measures of fatigue (to evaluate test burden) and malinger are planned.

Where feasible, test-retest reliability for several of the tasks is being assessed during current task evaluation trials at UNC and SKRC (Winter/Spring 2012) (**NOTE: DELETED**; this was written into the original statement of work (SOW) however, it has been determined that it is not feasible to complete

formal, reportable testing for test-retest reliability during the current grant work. Retest reliability for the multitasks would require parallel forms of the tasks and this should be addressed in future validation trials.

### Progress toward Aim 2:

In addition to the previous work (HC subjects at USARIEM) reported in the Annual Report for YEAR I (submitted September 2013), we have conducted additional IRR testing with subjects at Fort Bragg in two phases (Smith et al., 2014).

- Phase I:** Fort Bragg/Womack Army Medical Center data collection was completed between August and November 2013 in support of further evaluating the 6 AMMP test tasks for IRR of their scoring metrics in subjects with mTBI with subsequent revisions. This subject group included 20 subjects with mTBI. Due to the difficulty of scheduling multiple subjects within a few days, not all subjects could be rated by all 3 raters at the same time, therefore IRR data was evaluated for a subset of 13 subjects with mTBI who were rated in person by all 3 raters (2 physical therapists and 1 occupational therapist, all members of the AMMP development team). The exception to this was the evaluation of IRR of the PATROL-exertion (previously called the SALUTE-exertion) task. Revisions to the PATROL-exertion were ongoing during the August to November 2013 time frame so that the analysis reported here involves the 7 subjects tested on the final version of the SALUTE-exertion task before additional revisions were made to a PATROL report format as described under AIM 1 above. See TABLE 3 A-F below.
- Phase II:** Ongoing data collection continues at Fort Bragg/Womack Army Medical Center in support of the correlational and between groups analysis for AIMS 3 and 4. This phase of work began with the subjects tested beginning in January 2014. Along with construct and between groups validity work, we are continuing to collect inter-rater reliability data between 2 raters (both physical therapists) for a minimum of 18 subjects with mTBI and 5 HC subjects to further evaluate reliability of all task metrics following the revisions made after Phase I reliability testing. Preliminary analyses of these data are presented in Table 3 A-F below. Our rationale for testing additional subjects with mTBI has to do with the number and range of errors made by these subjects who provide for greater challenge in scoring their responses to the AMMP test tasks. Final analysis of IRR will be completed when these data are complete. Any metrics that do not meet expected reliability standards (See Table 2 above) will be deleted from the final between groups analysis.

**Table 3 (A-F): Inter-rater reliability (IRR) findings; Fort Bragg/Womack Army Medical Center**

| <b>A. CHARGE OF QUARTERS (CQ) DUTY</b>  |   |                                  |   |                                  |
|---|---|----------------------------------|---|----------------------------------|
| <b>Scoring item<br/>(Metrics)</b>       | <b>Fort Bragg/WAMC (Aug-Nov 2013)<br/>n=13 SM with mTBI</b> |                                  | <b>Fort Bragg/WAMC (Jan-Aug 2014) n= 21<br/>(14 mTBI, 7 HC)</b> |                                  |
|   | <b>Reliability<br/>ICC</b>                                  | <b>95% CI<br/>(lower, upper)</b> | <b>Reliability<br/>ICC</b>                                      | <b>95% CI<br/>(lower, upper)</b> |
| <b>Task<br/>performance</b>             | <b>0.90</b>   | <b>0.84-0.95</b>                 | <b>0.86</b>   | <b>0.70-0.97</b>                 |
| <b># of Rule<br/>breaks</b>             | <b>0.35</b>   | <b>0.03-0.62</b>                 | <b>0.83</b>   | <b>0.70-0.94</b>                 |
| <b># of Visits</b>                      | <b>0.92</b>   | <b>0.80-0.99</b>                 | <b>0.98*</b>  | <b>0.97-1*</b>                   |
| <b>Total time</b>                       | <b>0.99</b>   | <b>0.99-1.0</b>                  | <b>0.996</b>  | <b>0.99-1</b>                    |
| <b>* subjects 113, 114 not included</b> |   |                                  |   |                                  |



| B. SALUTE-EXERTION Version 2   |   |                          | PATROL-EXERTION                   |  |                          |
|--|---|--------------------------|-----------------------------------|--|--------------------------|
| Scoring item<br>(Metrics)  | Fort Bragg (Aug-Nov 2013) n=7<br>SM with mTBI |                          | Scoring item<br>(Metrics)         | Fort Bragg (Jan-Aug 2014) n=21 (14 mTBI, 7 HC) |                          |
|  | Reliability<br>ICC                            | 95% CI<br>(lower, upper) |                                   | Reliability<br>ICC                             | 95% CI<br>(lower, upper) |
| Total Sierra   | 3 triplets disagreed <sup>1</sup>             | NA                       | X. Sum A-D (IED marker reports)   | 0.96   | 0.92-0.99                |
| Total Alpha  | 5 triplets disagreed <sup>1</sup>             | NA                       | Y. Sum Post-test Patrol questions | 0.98   | 0.95-1                   |
| Total Tango  | 1 triplet disagreed <sup>1</sup>              | NA                       | Z. Sum of X and Y                 | 0.98   | 0.96-0.99                |
| Total Equipment  | 3 triplets disagreed <sup>1</sup>             | NA                       | Vision clarity pre-test           | 0.99   | 0.97-1                   |
| Scan IED Markers   | 0.97  | 0.94-0.99                | Vision clarity end                | 1.0  | 1-1                      |
| <ul style="list-style-type: none"> <li>After initial mTBI testing, further video revisions, removal of SALUTE format to “general Patrol reporting”, improved score sheets and instructions, addition of a auditory reaction time, and renamed PATROL-Exertion Task for future testing</li> </ul> <sup>1</sup> Given the small number of participants , we report number of triplets that disagreed |   |                          | Work RPE pre-test                 | 0.99   | 0.97-1                   |
|  |   |                          | Work RPE end                      | 1.0  | 1-1                      |

| C. RUN-ROLL-AIM  |  |                          |   |                          |
|--|--|--------------------------|---|--------------------------|
| Scoring item<br>(Metrics)  | Fort Bragg/WAMC (Aug-Nov 2013)<br>n=13 SM with mTBI                            |                          | Fort Bragg/WAMC (Jan-Aug 2014)<br>n= 21 (14 mTBI, 7 HC) |                          |
|  | Reliability<br>ICC   | 95% CI<br>(lower, upper) | Reliability<br>ICC                                      | 95% CI<br>(lower, upper) |
| Trial 1-Time(secs)   | 7 of 7-within 2 sec  | NA                       | 0.98  | 0.94-1                   |
| Trial 1-numbers correct  | 0.54   | 0.08-0.89                | 0.97  | 0.92-1                   |
| Trial 2-Time (secs)  | 7 of 7-within 1 sec  | NA                       | 0.997   | 0.995-0.999              |
| Trial 2-numbers correct  | 0.55   | 0.0-0.93                 | 0.996   | 0-1                      |
| Trial 3-Time(secs)   | 6 of 6*-within 1 sec   | NA                       | 0.995   | 0.99-1                   |
| Trial 3-numbers correct  | 0.72   | 0.40-0.95                | 0.996   | 0-1                      |
| Trial 4-Time(secs)   | 6 of 6*-within 2 sec   | NA                       | 0.999   | 0.998-0.999              |
| Trial 4-numbers correct  | 0.99   | 0.97-1.0                 | 0.98  | 0.96-1                   |
| Total errors (all trials)  | ICC's for individual trials calculated, T1: 0.54, T2: 0.13, T3: 0.18, T4: 0.85 |                          | 0.41  | 0-0.77                   |
| Total cues (all trials)  | Not measured   | NA                       | 0.83  | 0.60-0.98                |
| *not all subjects were able to tolerate completion of all trials<br>NA=not applicable or not evaluated |  |                          |   |                          |

| <b>D. LOAD MAGAZINE-RADIO CHATTER</b> |  |                          |  |                          |
|---------------------------------------|--|--------------------------|--|--------------------------|
| Scoring item<br>(Metrics)             | Fort Bragg/WAMC (Aug-Nov 2013) n=13 SM with mTBI |                          | Fort Bragg/WAMC (Jan-Aug 2014) n= 21 (14 mTBI, 7 HC) |                          |
|                                       | Reliability<br>ICC                               | 95% CI<br>(lower, upper) | Reliability<br>ICC                                   | 95% CI<br>(lower, upper) |
| Correct Key Word Single               | 0.94   | 0.88-0.99                | 0.96   | 0.89-1                   |
| Distractor Key Word Single            | 0.69   | 0.38-0.92                | 0.99   | 0.97-1                   |
| Correct Key Word Dual                 | 0.99   | 0.97-1.0                 | 0.99   | 0.98-1                   |
| Distractor Key Word Dual              | 0.50   | 0.11-0.82                | 0.98   | 0.94-1                   |

| <b>E. ILLINOIS AGILITY-WORD LIST</b>  |  |                          |  |                          |
|---|--|--------------------------|--|--------------------------|
| Scoring item<br>(Metrics)   | Fort Bragg/WAMC (Aug-Nov 2013) n=13 SM with mTBI |                          | Fort Bragg/WAMC (Jan-Aug 2014) n= 18 (13 mTBI, 5 HC) |                          |
|   | Reliability<br>ICC                               | 95% CI<br>(lower, upper) | Reliability<br>ICC                                   | 95% CI<br>(lower, upper) |
| Single Task Time  | 1.0  | NA                       | 0.99   | 0.98-0.99                |
| Single Task Words Correct   | 0.80   | 0.69-0.90                | 1.0  | 1-1                      |
| Single Task Word Errors   | 0.54   | 0.12-0.83                | 1.0  | 1-1                      |
| Dual Task No Instruction: Time  | 1.0  | NA                       | 0.99   | 0.98-0.99                |
| Dual Task NI: Words Correct   | 0.93   | 0.86-0.99                | 0.98   | 0.95-1                   |
| Dual Task NI: Word Errors   | 0.93   | 0.87-0.97                | 0.97   | 0.92-1                   |
| Dual Task NI: Course Errors   | NA   | NA                       | 1.0  | 1-1                      |
| Dual Task COG: Time   | 1.0  | 1-1                      | 0.99   | 0.99-1                   |
| Dual Task COG: Words Correct  | 0.97   | 0.92-1                   | 1.0  | 1-1                      |
| Dual Task COG: Word Errors  | 0.74   | 0.37-0.99                | 0.98   | 0.95-1                   |
| Dual Task COG: Course Errors  | NA   | NA                       | 0.78   | 0.70-1                   |
| Dual Task MOB: Time   | 1.0  | 1-1                      | 0.99   | 0.97-0.99                |
| Dual Task MOB: Words Correct  | 1.0  | 1-1                      | 1.0  | 1-1                      |
| Dual Task MOB: Errors   | 0.85   | 0.64-1                   | 0.96   | 0.88-1                   |
| Dual Task MOB: Course Errors  | NA   | NA                       | 1.0  | 1-1                      |
| <b>COG:</b> Cognitive priority; “concentrate on remembering the words”,<br><b>NI:</b> no instruction given,<br><b>MOB:</b> Mobility priority; “concentrate on going as fast as you can”<br><b>NA:</b> not applicable or not evaluated |  |                          |  |                          |

| <b>F. INSTRUMENTED STAND AND WALK-GRID COORDINATES (ISAW)</b> |   |                                  |  |                                  |
|---|---|----------------------------------|--|----------------------------------|
| <b>Scoring item<br/>(Metrics)</b>                             | <b>Fort Bragg/WAMC (Aug-Nov 2013)<br/>n=13 SM with mTBI</b> |                                  | <b>Fort Bragg/WAMC (Jan-Aug 2014) n=21 (14 mTBI, 7 HC)</b> |                                  |
|   | <b>Reliability<br/>ICC</b>                                  | <b>95% CI<br/>(lower, upper)</b> | <b>Reliability<br/>ICC</b>                                 | <b>95% CI<br/>(lower, upper)</b> |
| <b>Walk Time 1 Single</b>                                     | 0.77  | 0.64-0.86                        | 0.97   | 0.95-0.98                        |
| <b>Walk Time 2 Single</b>                                     | 0.95  | 0.92-0.98                        | NA   | NA                               |
| <b>Walk Time 3 Single</b>                                     | 0.91  | 0.85-0.96                        | NA   | NA                               |
| <b>Walk Time 1 Dual</b>                                       | 0.89  | 0.78-0.98                        | 0.91   | 0.82-0.97                        |
| <b>Walk Time 2 Dual</b>                                       | 0.94  | 0.89-0.96                        | 0.94   | 0.88-0.98                        |
| <b>Walk Time 3 Dual</b>                                       | 0.81  | 0.72-0.88                        | 0.98   | 0.96-0.99                        |
| <b>Grid Coord Single</b>                                      | 0.88  | 0.78-0.97                        | 0.997  | 0.991-1                          |
| <b>Grid Coord 1 Dual</b>                                      | 0.94  | 0.85-1                           | 0.95   | 0.86-1                           |
| <b>Grid Coord 2 Dual</b>                                      | 0.99  | 0.99-1                           | 0.998  | 0.995-1                          |
| <b>Grid Coord 3 Dual</b>                                      | 1.0   | 1-1                              | 0.99   | 0.98-1                           |
| <b>NA: not applicable or not evaluated</b>                    |   |                                  |  |                                  |

### Inter-rater reliability (IRR)

- The Krippendorff Alpha (Hayes 2007) was used to evaluate inter-rater reliability. This general measure can be used regardless of the number of observers, sample size, missing data and type of measurement (nominal, ordinal, interval, or ratio). For both interval and ratio data the analysis is equivalent to the intraclass correlation coefficient (ICC) for two observers and is extended for many observers. For nominal data, analysis for two observers is equivalent to Scott's Pi. Parallel analyses using both the Krippendorff and Kappa (2 observers) have produced identical results. The code was integrated into SPSS V18.0. Bootstrapping using an n=2000 was used to produce 95% confidence intervals.
- For low ICCs further analysis of paired scorers was completed to discover reasons for reliability errors (not shown in this report) during the Aug-Nov 2013 time frame.

### Limitations

Inter-rater reliability calculations are highly sensitive to the range of possible values. For items that take on values such as 0 to 5, differences between scorers will affect the calculation greater than values that are continuous such as time or total number of tasks completed that have high maximum values. This is due to the distance between values relative to the range.

Due to this sensitivity, sub-task groupings as is used for several multitasks may have low IRR while the total IRR is acceptable. This supports the need to evaluate sub-group scores and perhaps even item by item evaluation may be necessary where a large number of observations are required. Between the 2 phases of IRR testing, modifications to some score sheets were made to eliminate items with low values for their range, as appeared reasonable.

### Implications:

The data presented here for January-August 2014 is an interim look at the findings of IRR data between two testers at Fort Bragg/WAMC which was initiated in January 2014 in concert with the data for AIMS 3 and 4 (correlation to neurocognitive tests and between groups discriminate validity). These data show good to excellent interrater reliability for most AMMP task metrics following the task instruction and

scoring revisions. A minimum of 18 subjects with mTBI and 5 HC subjects will be evaluated before final IRR analysis is completed. We recognize that the testing on subjects with mTBI has expanded the range of scores on test tasks and the type of responses and errors seen on the task metrics.

Overall, Aims 1 and 2 for this protocol at Fort Bragg/Womack Army Medical Center (August 2013 to date) are being met. Refining the test burden down to two hours is one of the goals for this phase of testing at Fort Bragg which began in August 2013. The AMMP team will continue to use the reliability and feasibility data collected in the Fort Bragg/WAMC protocol to make final determinations regarding test battery composition. Additional reliability testing will be recommended for future studies.

**Aim 3: Determine correlation between scores on neurobehavioral and sensorimotor domain tests and scores on AMMP dual- and multitasks in healthy control SMs and SM with mTBI.**

Approach: Preliminary evaluation of construct validity will be performed by determining correlations between neurobehavioral and sensorimotor tests of known mTBI vulnerabilities (e.g., clinical tests of selective attention, processing speed, working memory, executive function and dynamic visual acuity and individual AMMP tasks in healthy control SM and SM with mTBI symptom complex. Testing is being conducted at Fort Bragg /Womack Army Medical Center with a goal of 80 SM in each group, (healthy control [HC] and mTBI).

**Progress toward Aim 3:** Currently data collection is ongoing at Fort Bragg (beginning August 2013) with a total of **60 subjects tested to date (36 mTBI and 24 HC)**. All subjects complete the intake form which includes demographic and PCL-C findings. As well, subjects are tested on the Dynamic Visual Acuity Test and the neurocognitive battery which includes: 1) simple reaction time, 2) Neuropsychological Assessment Battery attention module, 3) Tower of Hanoi, 4) Comprehensive Trail Mailing Test, 5) Wide Range Achievement Test Reading, and 6) Test of Memory Malingering. Correlations between the scores on these standard tests and the AMMP test tasks metrics will be carried out after all data are collected.

**Aim 4: Determine ability of dual-task and multitask test items to discriminate between healthy control SM and SM with mTBI symptom complex.**

Approach: A known groups comparison will be used to evaluate the ability of individual AMMP tasks to discriminate between 80 HC SM and 80 SM with diagnosed mTBI symptom complex. Testing is being completed in concert with Aim 3 at Fort Bragg/Womack Army Medical Center. These data will be evaluated with an ANOVA. Estimates of anticipated effect sizes for determining sample size have been determined from available literature and drive our hypothesized minimum effect size of 0.5.

**Progress toward Aim 4:** Studies at Fort Bragg/Womack Army Medical Center in Fayetteville, NC received all regulatory approvals in August 2013 and to date a total of 60 subjects (36 mTBI and 24 HC) have been tested (see Progress description under Aim 3). Subjects are scheduled as available on a weekly basis with data collection anticipated to continue through approximately 1 April 2015. Analysis of all metrics for AMMP test tasks including inertial sensor variables will be fully evaluated at the completion of data collection.

Interim findings for the Instrumented Stand and Walk (ISAW) test have undergone preliminary evaluation. The ISAW test (Mobility Lab, APDM INC, Portland OR) uses small, wireless Opal™ movement monitors which contain 3D angular rate sensor, 3D accelerometer and gyroscope. Monitors were affixed to the lumbar area and lateral ankles to quantify sway, gait, and rotational kinematics. A commercial algorithm is used to analyze all movements to include the initiation and completion of turns. The interim evaluation used a convenience sample of 34 healthy (23 male, age  $26.9 \pm 5.1$  years) Soldiers and 30 active duty (all male; age  $28.7 \pm 6.7$  years) Soldiers receiving treatment for persistent post-concussive

symptoms at a military TBI care center. Median values from three-replication sets were compared between groups for the sway and gait parameters using a two-sided t-test; p-values <0.05 are reported. Measures of sway during quiet stance and gait at a self-selected, "comfortable" walking pace were not significantly different between groups. Instrumented measures of turning revealed significant between group differences. Soldiers with post-concussive deficits demonstrated longer turn durations (HC:  $1.58 \pm 0.29$  seconds; mTBI:  $1.90 \pm 0.37$  seconds,  $p < .001$ ); increased step numbers to complete a turn (HC:  $3.51 \pm 0.56$ ; mTBI:  $4.07 \pm 0.87$  steps,  $p < .004$ ), and decreased peak rotational velocities during turns (HC:  $225.73 \pm 45.49^\circ/\text{s}$ ; mTBI:  $192.10 \pm 35.33^\circ/\text{s}$ ,  $p < .003$ ). Based on this preliminary analysis, instrumented assessment with inertial sensors shows promise as a means to detect subtle post-concussive differences between the healthy control and subjects with concussion undergoing treatment for persistent deficits. Again, final evaluation of the ISAW clinical and inertial sensor data in both single and dual-task conditions (cognitive task requires memorization of grid coordinates) will be done following completion of all data collection.

Work is progressing on the analysis techniques for the inertial sensor data (NexGen Ergonomics-movement monitors similar to the wireless Opal™ system) from the head and waist sensors that are used during the Run-Roll-Aim (RRA) Task and the Illinois Agility-Word List Dual-task (IAT). This requires development of appropriate software as no commercially obtainable programs are available for analysis of these unique test tasks. Currently we are exploring the use of non-linear methods including cepstrum coefficients, deriving displacement data from acceleration; and phase-space plots in order to evaluate which components of the movements in these two tasks show promise in distinguishing the subject groups.

Using the accelerometer data recorded by the inertial sensors, we can plot a representation of the body movements in what is called a "phase space". The phase space has axes of head or torso acceleration plotted against the same acceleration data only with a forward time delay. This visually represents the relationship between the acceleration of the body at a given point in time and its acceleration at a given time in the future. From this plot the predictability of the movement being analyzed can be estimated. We are exploring the use of the Maximal Lyapunov exponent (MLE) as a means to quantitatively measure the predictability of body movement. A positive MLE is usually taken as an indication that the system is chaotic. Therefore we hypothesize that HC subjects will have a lower MLE, indicating more predictability which translates to more dynamic stability, and BI subjects will have a higher MLE, indicating less predictability and less dynamic stability. Currently we are evaluating the entire Illinois Agility Test trace for the single task trial using the above described methods. The decision on whether to analyze the entire Run-Roll-Aim trace or to segment the data into its different components (i.e., the forward run, roll, the stationary target search, side shuttle or backpedal) is being discussed.

#### **KEY RESEARCH ACCOMPLISHMENTS:**

- Continuing Review approvals of our protocol have been received from Womack Army Medical Center IRB and Quorum Health IRB for this project as of 12 May 2014. (NOTE: All Allina Health dual-oversight IRB protocols have been transferred to Quorum Health IRB beginning January 2014).
- A standardized AMMP administration manual is close to complete. This includes administration instructions for all 6 test tasks and administration instructions for the neurocognitive tests and dynamic visual acuity test being administered in this protocol. Preliminary training videos on

AMMP tasks are being developed for use in future clinical and validation studies. This video material will be available in draft form in early 2015.

- Currently data collection is ongoing at Fort Bragg (beginning August 2013) with a total of 60 subjects tested to date (36 mTBI and 24 HC). A total of 587 subjects have been briefed and a total of 163 subjects who initially expressed interest have been contacted. Cancellations due to command requests, weather related closures and volunteer subjects deciding not to participate have limited subject numbers. A request for a one year (through 14 August 2015) no cost extension was approved by USAMRAA on 4April 2014 to allow continuing data collection in support of the AIMs of this project.

## **REPORTABLE OUTCOMES:**

### **1) Papers**

- a. Smith, LB, Radomski, MV, Davidson LF, Finkelstein, M, Weightman, MM, Scherer, MR, McCulloch, K (2014). Development and preliminary reliability of a multitasking assessment following concussion. *American Journal of Occupational Therapy*, 68, 439-443.

### **2) Presentations (Symposium, Platform or Poster)**

- a. "Returning Service Members to Duty Following Mild Traumatic Brain Injury: Exploring the Use of Dual- and Multitask Assessment Methods", Karen L. McCulloch PT, PhD, NCS, University of North Carolina, Chapel Hill, NC, AMSUS Conference (The Society of Federal Health Professionals), 6 November 2013, Seattle, WA.
- b. "Assessing Motor Performance in Military Service Members Following mTBI Using Inertial Sensor Data, Yu R, McCulloch K, Human Movement Science Student Research Day, University of North Carolina, Chapel Hill, NC, 21 February 2014.
- c. "Preliminary inter-rater reliability for a novel dual-task and multitask assessment battery guiding return-to-duty in concussed Service Members", Margaret Weightman, Karen McCulloch, Leslie Freeman Davidson, Matthew Scherer, Laurel Smith, Marsha Finkelstein, Mary Vining Radomski, International Brain Injury Association Conference, 20March 2014, San Francisco, CA.
- d. "Development of a Dual-task and Multitask Assessment to inform Return-to-Duty after Concussion: Implications for Civilian Practice", Weightman MM, Radomski MV, MN Brain Injury Association Annual Meeting, 11April 2014, Roseville, MN.
- e. "State of the Art Military TBI Care: Latest Policies and Innovations in Return-to-Duty Assessment", McCulloch K, Maxfield-Panker S, McMillan H, American Physical Therapy Association's Annual Conference and Exposition, 12June 2014, Charlotte, NC.
- f. "A novel dual-task and multitask assessment battery guiding return-to-duty in concussed Service Members", Laurel Smith, Mary Vining Radomski, Marsha Finkelstein, Karen McCulloch, Leslie Freeman Davidson, Henry McMillan, Matthew Scherer, Margaret Weightman, 3<sup>rd</sup> International Congress on Soldiers' Physical Performance, 20August 2014 Boston, MA.

### **3) Abstracts Accepted for poster or platform presentations**

- a. Poster #2055 "Development of the "Run-Roll-Aim" Task, A High Level Mobility Task Assessing Physical Skills Related to Military Duty"; Poster # 2053 "Development of the Dual-task Illinois Agility Task, A Component of the Assessment of Military

Multitask Performance”, American College of Rehabilitation Medicine Annual Conference, McCulloch K, Goldberg R, 10 October 2014, Toronto, Ontario, CA.

- b. American Physical Therapy Association’s Combined Sections Meeting, “Inertial sensors detect subtle mobility differences in soldiers with persistent concussion symptoms: preliminary findings for the instrumented stand and walk”, M Scherer, M Finkelstein, K McCulloch, L Smith, M Weightman, 4-7 February 2015 Indianapolis, Indiana.

#### 4) Works in preparation or submitted ( awaiting acceptance):

- Manuscript in preparation for Journal of *Rehabilitation Research & Development*, anticipated initial submission December 2014 and tentative title: “Iterative Reliability Testing for Revision of Ecologically Valid Performance-Based Test Tasks: Refinement of the Assessment of Military Multitasking Performance” Weightman M, Radomski MV, Finkelstein M, Smith L, Davidson L, McCulloch K, Scherer M.

#### CONCLUSION:

A research team of military and civilian physical and occupational therapists is working to refine and develop preliminary validation data on a set of novel test-tasks which are part of the Assessment of Military Multitasking Performance (AMMP). Once further validated, the AMMP is anticipated to be used in combination with other metrics to inform duty-readiness decisions for Service members following mTBI. Previous inter-rater reliability findings from healthy control Service member volunteers from USARIEM have been expanded upon in testing healthy and concussed Service Members from Fort Bragg/Womack Army Medical Center. Achieving adequate IRR on ecologically valid dual-tasks and multitasks requires the data-driven iterative process of refinement of metrics, procedures, and task properties.

Studies at Fort Bragg/Womack Army Medical Center in Fayetteville, NC received all regulatory approvals in August 2013 and to date a total of 60 subjects (36 mTBI and 24 HC) have been tested in support of the known groups and correlational validation study aims. The AMMP test tasks currently demonstrate good to excellent inter-reliability for most test metrics (See Table 3A-F). Task modifications were made based on test burden and inter-rater reliability findings for AMMP multitasks; further reliability testing on SM with mTBI is ongoing at Fort Bragg. Data collection is ongoing with a goal of 80 healthy control and 80 subjects with mTBI symptom complex. Data analysis is now expected to be complete in early summer 2015 following approval of a no cost extension through August 2015. Dissemination efforts on the progress and findings of this work are ongoing.

#### REFERENCES:

- 1) Radomski MV, Weightman MM, Davidson LF, Finkelstein M, Goldman S, McCulloch KL, Roy TC, Scherer M, Stern EB,. Development of a Measure to Inform Return-to-Duty Decision Making after Mild Traumatic Brain Injury. *Military Medicine*, 178(3):246-253, 2013.
- 2) Matthew R. Scherer, Margaret M. Weightman, Mary V. Radomski, Leslie F. Davidson, Karen L. McCulloch; Returning Service Members to Duty Following Mild Traumatic Brain Injury: Exploring the Use of Dual-Task and Multitask Assessment Methods. *Physical Therapy Journal*, Sept 2013, 93(9):1-14.

- 3) Behavioral Assessment of the Dysexecutive Syndrome, Wilson, B.A., Alderman, N., Burgess, P.W., Hazel, E., Evans, J.J. (1996) Pearson Assessment.
- 4) Smith, LB, Radomski, MV, Davidson LF, Finkelstein, M, Weightman, MM, Scherer, MR, McCulloch, K (2014). Development and preliminary reliability of a multitasking assessment following concussion. *American Journal of Occupational Therapy*, 68, 439-443.
- 5) Hayes AF, Krippendorff K, (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1:77-89.

#### **APPENDICES:**

- 1) (ABSTRACT) 3<sup>rd</sup> International Congress on Soldiers' Physical Performance, "A novel dual-task and multitask assessment battery guiding return-to-duty in concussed Service Members", Laurel Smith, Mary Vining Radomski, Marsha Finkelstein, Karen McCulloch, Leslie Freeman Davidson, Henry McMillan, Matthew Scherer, Margaret Weightman, 20August 2014 Boston, MA.
- 2) (ABSTRACT) International Brain Injury Association Conference, "Preliminary inter-rater reliability for a novel dual-task and multitask assessment battery guiding return-to-duty in concussed Service Members", Margaret Weightman, Karen McCulloch, Leslie Freeman Davidson, Matthew Scherer, Laurel Smith, Marsha Finkelstein, Mary Vining Radomski, 20March 2014, San Francisco, CA.
- 3) (ABSTRACT-accepted) American Physical Therapy Association's Combined Sections Meeting, "Inertial sensors detect subtle mobility differences in soldiers with persistent concussion symptoms: preliminary findings for the instrumented stand and walk", M Scherer, M Finkelstein, K McCulloch, L Smith, M Weightman, 4-7 February 2015 Indianapolis, Indiana.
- 4) Smith, LB, Radomski, MV, Davidson LF, Finkelstein, M, Weightman, MM, Scherer, MR, McCulloch, K (2014). Development and preliminary reliability of a multitasking assessment following concussion. *American Journal of Occupational Therapy*, 68, 439-443.



### **3<sup>rd</sup> International Congress Soldier Physical Performance, 20 August 2014, Boston MA**

**Title:** A novel dual-task and multitask assessment battery guiding return-to-duty in concussed Service Members.

Laurel Smith<sup>1</sup>, Mary Vining Radomski<sup>2</sup>, Marsha Finkelstein<sup>2</sup>, Karen McCulloch<sup>3</sup>, Leslie Freeman Davidson<sup>4</sup>, Henry McMillan<sup>5</sup>, Matthew Scherer<sup>6</sup>, Margaret Weightman<sup>2</sup>

1 United States Army Research Institute of Environmental Medicine, Natick, MA

2 Courage Kenny Research Center, Allina Health, Minneapolis, MN

3 Division of Physical Therapy, Department of Allied Health Sciences, University of North Carolina, Chapel Hill, NC

4 Riverbend Therapeutics, LLC, Great Falls, VA

5 Department of Brain Injury Medicine, Womack Army Medical Center, Fort Bragg, Fayetteville, NC

6 Andrew Rader US Army Health Clinic, Joint Base Myer-Henderson Hall, Fort Myer, VA

**PURPOSE:** The Assessment of Military Multitasking Performance (AMMP) is a battery of military-related functional dual-tasks and multitasks that target known sensorimotor, cognitive, and exertional vulnerabilities after concussion/mild traumatic brain injury (mTBI). Once validated, the AMMP is intended for use in combination with other metrics to inform duty-readiness decisions in service members following concussion. A dual task paradigm requires a Soldier to perform a physical and a cognitive task simultaneously in order to compare dual-task with single task performance. A multitask format requires completion of complex physical and cognitive activities that approximate real-world military tasks. Several test tasks challenge agility or activity tolerance. Initial validation for the AMMP involves establishing interrater reliability (IRR); and then convergent/discriminant validity by using correlations to neurocognitive and sensorimotor tests and establishing known groups validity by comparing scores on AMMP tasks between healthy control (HC) and concussed Soldiers undergoing rehabilitation in a brain injury clinic. **METHODS:** Using a convenience sample case-control methodology involving test construction and evaluation, a data-driven iterative process has been used to evaluate the six AMMP test tasks for interrater reliability (IRR) by 3 person rater teams comprised of physical and occupational therapists. Scoring discrepancies identified by intraclass correlation coefficients resulted in further clarifications of scoring rules and scorer training requirements. Ongoing data collection efforts continue at Fort Bragg for both HC and concussed Soldiers with a goal of 80 subjects per group. **RESULTS:** In addition to preliminary HC reliability testing, 34 subjects have been tested to date. Reliability findings frequently differed in HC versus concussed groups. ICCs for task completion time were 0.96-0.99 in HC and 0.77 to 0.99 in subjects with concussion. Cognitive components for each of the 3 dual-tasks, such as responding to key words in recorded radio chatter or recalling grid coordinates, demonstrated ICCs between 0.64 and 0.99. Subjects with concussion typically demonstrated greater number and range of errors than were seen in HC. **CONCLUSIONS:** Preliminary testing informed modifications in test structure, instruction, and scoring to enhance IRR. Development of measures that meet military stakeholder requirements for face validity and functional relevance contribute to the challenges of development of a valid AMMP battery. The consistency of scores across raters and the

ability to discriminate known groups are fundamental to using the findings of the AMMP to make substantive recommendations regarding readiness to return to duty following concussion/mTBI.

Funding for this work provided by MPMC W81XWH-12-2-0070.

### **International Brain Injury Association Meeting Abstract, 20 March 2014 San Francisco, CA**

**Title:** Preliminary inter-rater reliability for a novel dual-task and multitask assessment battery guiding return-to-duty in concussed Service Members.

Margaret Weightman<sup>1</sup>, Karen McCulloch<sup>2</sup>, Leslie Freeman Davidson<sup>3</sup>, Matthew Scherer<sup>4</sup>, Laurel Smith<sup>5</sup>, Marsha Finkelstein<sup>1</sup>, Mary Vining Radomski<sup>1</sup>

1 Courage Kenny Research Center, Allina Health, Minneapolis, MN

2 Division of Physical Therapy, Department of Allied Health Sciences, School of Medicine, University of North Carolina, Chapel Hill, NC

3 Riverbend Therapeutics, LLC, Great Falls, VA

4 Andrew Rader US Army Health Clinic, Joint Base Myer-Henderson Hall, Fort Myer, VA

5 United States Army Research Institute of Environmental Medicine, Natick, MA

**OBJECTIVES:** The Assessment of Military Multitasking Performance (AMMP) is a battery of military-related functional dual-tasks and multitasks that target known sensorimotor, cognitive, and exertional vulnerabilities after concussion/mild traumatic brain injury (mTBI). Once validated, the AMMP is intended for use in combination with other metrics to inform duty-readiness decisions in active duty service members following concussion. The assessment of inter-rater reliability (IRR) provided data regarding IRR of individual AMMP tasks, and informed training requirements needed for reliable scoring of the battery. Preliminary IRR findings highlight the challenges and successes in development of performance based assessments designed to identify subtle deficits in highly trained personnel.

**METHODS:** Six AMMP test tasks were individually evaluated for IRR by 3 person rater teams comprised of physical and occupational therapists who were the task developers, with at least one rater who was initially unfamiliar with the task. Initial IRR evaluation for 4 tasks was completed on 20 healthy Soldiers (HC), and for 2 tasks on 12 HC. Scoring discrepancies identified by the statistical analysis using Krippendorff Alpha resulted in further clarifications of scoring rules and scorer training requirements. Tasks were again tested by 3 raters on 11 to 13 Soldiers undergoing rehabilitation following concussion/mTBI. **RESULTS:** Reliability findings frequently differed in HC versus concussed groups. For example, ICCs for task completion time were 0.96-0.99 in HC and 0.77 to 0.99 in subjects with concussion. Cognitive components for each of the 3 dual-tasks, such as responding to key words in recorded radio chatter or recalling grid coordinates, demonstrated ICCs between 0.64 and 0.99. Multitask metrics demonstrated variable ICCs. For example, task completion and number of transits

during a 'Charge of Quarters' multitask demonstrated excellent IRR (ICCs of 0.90 to 0.98). IRR calculations were highly sensitive to the range of possible values with metrics that involve restricted ranges such as number of errors, cues, or rule breaks, demonstrating variable and often lower ICCs (ICC range 0.13-0.85). Subjects with concussion typically demonstrated greater number and range of errors than were not seen in testing the healthy control Soldiers. **CONCLUSIONS:** Preliminary IRR testing informed modifications in test instruction, structure, and scoring to enhance IRR. Development of measures that meet military stakeholder requirements for face validity and functional relevance contribute to the complexity of development of a reliable AMMP battery. The consistency of scores across raters is fundamental to the ability to use the findings of the AMMP to make substantive recommendations regarding readiness to return to duty following concussion/mTBI.

Funding for this work provided by MRMC W81XWH-12-2-0070. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Army or the Department of Defense.

**American Physical Therapy Association Combined Sections Meeting Abstract, 4-7February 2015, Indianapolis, IN**

Authors: M Scherer<sup>1</sup>, M Finkelstein<sup>2</sup>, K McCulloch<sup>3</sup>, L Smith<sup>4</sup>, M Weightman<sup>2</sup>

Affiliations: 1 Physical Therapy Service, Andrew Rader US Army Health Clinic, Joint Base Ft. Myer-Henderson Hall, VA

2 Courage Kenny Research Center, Minneapolis, MN

3 Division of Physical Therapy, University of North Carolina at Chapel Hill, Chapel Hill, NC

4 US Army Research Institute of Environmental Medicine, Natick, MA

**Title: Inertial sensors detect subtle mobility differences in soldiers with persistent concussion symptoms: preliminary findings for the instrumented stand and walk**

**Purpose/Hypothesis:** Imbalance following concussion often results from abnormal sensory integration between visual, somatosensory and vestibular inputs. Subtle postural control and mobility deficits may impact performance, safety, and readiness to return to duty (RTD) or play following concussion/mild traumatic brain injury (mTBI) and go undetected by standard clinical balance and gait tests. The purpose of this report is to describe findings from an instrumented 30 second stand and 7 meter walk with a 180 degree turn for the ability to distinguish healthy control (HC), "duty ready" Soldiers from Soldiers undergoing rehabilitation for persistent deficits following mTBI.

**Subjects:** A convenience sample of 34 healthy (23 male, age  $26.9 \pm 5.1$  years) Soldiers and 30 active duty (all male; age  $28.7 \pm 6.7$  years) Soldiers receiving treatment for persistent post-concussive symptoms at a military TBI care center.

**Materials/Methods:** The Instrumented Stand and Walk (ISAW) test (Mobility Lab, APDM INC, Portland OR) uses small, wireless Opal™ movement monitors which contain 3D angular rate sensor, 3D accelerometer and gyroscope. Monitors were affixed to the lumbar area and lateral ankles to quantify sway, gait, and rotational kinematics. A commercial algorithm was used to analyze all movements to include the initiation and completion of turns. Three trials of the ISAW were completed. Median values from three-replication sets were compared between groups for the sway and gait parameters using a two-sided t-test; p-values <0.05 are reported.

**Results:** Measures of sway during quiet stance and gait at a self-selected, “comfortable” walking pace were not significantly different between groups. Instrumented measures of turning revealed significant between group differences. Soldiers with post-concussive deficits demonstrated longer turn durations (HC:  $1.58 \pm 0.29$  seconds; mTBI:  $1.90 \pm 0.37$  seconds,  $p < .001$ ); increased step numbers to complete a turn (HC:  $3.51 \pm 0.56$ ; mTBI:  $4.07 \pm 0.87$  steps,  $p < .004$ ), and decreased peak rotational velocities during turns (HC:  $225.73 \pm 45.49^\circ/s$ ; mTBI:  $192.10 \pm 35.33^\circ/s$ ,  $p < .003$ ).

**Conclusions:** These findings support the utility of the ISAW to detect subtle sensorimotor deficits in Soldiers with persistent post-concussive deficits during a 180 degree turn which involved rapid peak rotational movements. Analysis of sway and gait kinematics did not reveal significant between groups differences.

**Clinical Relevance:** Based on this preliminary analysis, instrumented assessment with inertial sensors shows promise as a means to detect subtle post-concussive differences where standard clinical measures of mobility and balance may be insufficient in highly trained military personnel. This type of instrumented assessment in combination with other functional metrics may more fully characterize readiness to RTD or play.

**Keywords:** concussion, inertial sensors, turning

## REFERENCES:

1. Scherer M, Weightman M, Radomski M, Davidson L, McCulloch K. “Returning Service Members to Duty Following Mild Traumatic Brain Injury: Exploring the Use of Dual-Task and Multitask Assessment Methods”, *Phys Ther.* 2013; 93:1255-1267.
2. Mancini M, Salarian A, Carson-Kuhta P, Zampieri , Chiari L., Horak F. I-SWAY: a sensitive, Valid and reliable measure of Postural control. *Trans Neural Systems Rehab Eng* 2010.
3. King LA, Mancini M, Priest K, Salarian A, Rodrigues-de-Paula, Horak F. Do clinical scales of balance reflect turning abnormalities in people with Parkinson’s disease”, *J Neurol Phys Therapy* 2012 March; 36(1):25-31.
4. Scherer M, Burrows H, Pinto R, Littlefield P, French L, Tarbett A, Schubert M. Evidence of Central and Peripheral Vestibular Pathology in Blast Related Traumatic Brain Injury”, *Otol Neurotol*, 2011 Jun;32(4):571-80.
5. Martini DN, Sabin MJ, DePesa SA, Leal EW, Negrete TN, Sosnoff JJ, Broglio SP. The chronic effects of concussion on gait. *Arch Phys Med Rehabil* 2011;92:585-9.
6. Salarian A, Zampieri C, Horak F, Carlson-Kuhta P, Nutt JG, Aminian K. Analyzing 180° turns using an inertial system reveals early signs of progress in Parkinson’s Disease. *Conf Proc IEEE Eng Med Biol Soc.* 2009; 224-227.
7. King, LA, Horak FB, Mancini M, Pierce D, Priest KC, Chesnutt J, Sullivan P, Chapman JC. Instrumenting the Balance Error Scoring System for use with patients reporting persistent balance problems after mild traumatic brain injury. *Arch Phys Med Rehab* 2014; 95(2):353-359.

# Development and Preliminary Reliability of a Multitasking Assessment for Executive Functioning After Concussion

Laurel B. Smith, Mary Vining Radomski, Leslie Freeman Davidson, Marsha Finkelstein, Margaret M. Weightman, Karen L. McCulloch, Matthew R. Scherer

## MeSH TERMS

- brain concussion
- executive function
- military personnel
- reproducibility of results
- task performance and analysis

**OBJECTIVES.** Executive functioning deficits may result from concussion. The Charge of Quarters (CQ) Duty Task is a multitask assessment designed to assess executive functioning in servicemembers after concussion. In this article, we discuss the rationale and process used in the development of the CQ Duty Task and present pilot data from the preliminary evaluation of interrater reliability (IRR).

**METHOD.** Three evaluators observed as 12 healthy participants performed the CQ Duty Task and measured performance using various metrics. Intraclass correlation coefficient (ICC) quantified IRR.

**RESULTS.** The ICC for task completion was .94. ICCs for other assessment metrics were variable.

**CONCLUSION.** Preliminary IRR data for the CQ Duty Task are encouraging, but further investigation is needed to improve IRR in some domains. Lessons learned in the development of the CQ Duty Task could benefit future test development efforts with populations other than the military.

Smith, L. B., Radomski, M. V., Davidson, L. F., Finkelstein, M., Weightman, M. M., McCulloch, K. L., & Scherer, M. R. (2014). Brief Report—Development and preliminary reliability of a multitasking assessment for executive functioning after concussion. *American Journal of Occupational Therapy*, 68, 439–443. <http://dx.doi.org/10.5014/ajot.2014.012393>

**Laurel B. Smith, MS, OTR/L**, is Captain, U.S. Army, and Research Occupational Therapist, U.S. Army Research Institute of Environmental Medicine, 15 Kansas Street, Natick, MA 01760; [laurel.b.smith.mil@mail.mil](mailto:laurel.b.smith.mil@mail.mil)

**Mary Vining Radomski, PhD, OTR/L**, is Clinical Scientist, Courage Kenny Research Center, Minneapolis, MN.

**Leslie Freeman Davidson, PhD, OTR/L**, is Director and Associate Professor of Occupational Therapy, Shenandoah University, Winchester, VA.

**Marsha Finkelstein, MS**, is Senior Scientific Advisor, Courage Kenny Research Center, Minneapolis, MN.

**Margaret M. Weightman, PhD, PT**, is Clinical Scientist/Physical Therapist, Courage Kenny Research Center, Minneapolis, MN.

**Karen L. McCulloch, PhD, PT, NCS**, is Professor, Division of Physical Therapy, University of North Carolina at Chapel Hill.

**Matthew R. Scherer, PhD, PT**, is NCS Major, U.S. Army, and Chief of Physical Therapy, Andrew Rader U.S. Army Health Clinic, Fort Myer, VA.

Concussion has received unprecedented attention in the military because of the increased incidence in the past decade (Helmick, Baugh, Lattimore, & Goldman, 2012) and has been called the “signature injury” of the conflicts in Iraq and Afghanistan (McCrea et al., 2009, p. 1369). Concussion may result in symptoms including headache, dizziness, nausea, sensitivity to noise and light, slowed thinking and reaction time, memory problems, difficulty concentrating, executive dysfunction, and visual and balance changes (Carroll et al., 2004). Although subtle and sometimes difficult to detect, these multisensory symptoms can negatively affect job performance and safety in servicemembers.

Army occupational therapists play key roles in evaluating servicemembers and making recommendations regarding their ability to return to duty after concussion. Currently, occupational therapy practitioners rely on self-reported symptoms and

vestibular and neuropsychological assessments to determine duty readiness. However, subjective symptom report does not always coincide with clinical recovery (Vagnozzi et al., 2008), and neuropsychological assessment batteries do not always predict real-world functioning, especially after a combat experience (Brenner et al., 2010). Accurate assessment is further limited by measures with ceiling effects or minimal sensitivity to concussion-related deficits.

Multitask assessments may be more sensitive to subtle performance deficits because they replicate the simultaneous cognitive and sensorimotor demands of unstructured, complex real-world activities (Frisch, Förstl, Legler, Schöpe, & Goebel, 2012). Despite the potential benefit of this assessment approach and alignment with priorities for occupational therapy evaluation, few options exist that have satisfactory reliability, validity, and clinical utility (Dawson et al., 2009). The

Multiple Errands Test (MET; Shallice & Burgess, 1991) is an example of a multi-task assessment of executive functioning based on five demands of multitasking: (1) performing multiple but discrete tasks that vary in priority, complexity, and length; (2) managing interleaving and dovetailing tasks; (3) performing tasks without feedback; (4) dealing with interruptions, reprioritization, and rule changes; and (5) self-initiating task changes within the activity (Burgess, 2000). The many versions of the MET involve completing at least 10 unrelated tasks while complying with a series of rules in either a shopping mall or hospital lobby setting (Alderman, Burgess, Knight, & Henman, 2003; Cuberos-Urbano et al., 2013; Dawson et al., 2009; Morrison et al., 2013). Although the MET appears to assess “the central aspects of executive functioning in everyday life” (Frisch et al., 2012, p. 257), it has yet to be widely adopted in clinical practice because of site-specific validation requirements, time-intensive administration, and a lack of standardized scoring manuals specific to each site (Radomski & Morrison, 2014).

A team of military and civilian occupational and physical therapists are currently developing a performance-based assessment battery called the Assessment of Military Multitasking Performance (AMMP; Radomski et al., 2013). The AMMP includes six dual- and multitask assessments designed to assess concussion-related deficits. If proven reliable and valid, the AMMP will be used by military occupational and physical therapists to determine duty readiness for servicemembers after concussion.

The Charge of Quarters (CQ) Duty Task (CQDT) was developed as one of the assessments included in the AMMP battery that uses the structure of the MET to assess executive functioning. CQ duty is an additional duty in the military during which servicemembers are responsible for 24-hr supervision and security of a facility; servicemembers on CQ duty are frequently tasked with various assignments that are unstructured and unrelated in nature. This scenario provides a realistic backdrop for the multitask assessment given the reality of task demands and

face validity among servicemembers. This article describes the rationale and development process of the CQDT and presents pilot data from the preliminary evaluation of interrater reliability (IRR).

## Description of the Charge of Quarters Duty Task

In the CQDT, as in the MET, participants receive in-depth instructions and a written list of assignments and performance rules. They are required to visit four different hypothetical work areas (marked with duct tape): (1) the CQ desk, (2) the bulletin board, (3) the supply closet, and (4) the assembly area, each containing the information and resources necessary to complete their assignments. They are encouraged to keep transits between work areas to a minimum (seven or fewer) and are told to revisit an area only if necessary to complete the task. Task assignments include reporting a CQ duty shift change, assembling a footstool from PVC pipe, reporting the number of vacant rooms in the barracks (living quarters for servicemembers) using a barracks layout, conducting an inventory of PVC supplies, obtaining the address of the manufacturer of the footstool materials, locating the telephone number of another servicemember using a personnel roster, and locating the room of a specified servicemember using a map of a barracks layout.

During the exercise, participants must adhere to four rules: (1) Assemble the footrest only in the assembly area, (2) bring only the number of PVC parts needed for the footrest to the assembly area, (3) do not move or remove any of the materials from the walls in any of the work areas, and (4) do not speak to the examiners during the assessment. Throughout the task, participants must also deal with interruptions and reprioritization of tasks. Scoring metrics borrowed from the MET include accuracy of task performance (Cuberos-Urbano et al., 2013; Dawson et al., 2009; Morrison et al., 2013), total rule breaks (Cuberos-Urbano et al., 2013; Dawson et al., 2009; Morrison et al., 2013), frequency of rule breaks (Dawson et al., 2009), transits between work areas (Morrison et al., 2013), and total performance time.

## Method

### *Instrument Development*

The CQDT was developed as part of the AMMP battery. The initial version of the AMMP included five multitask assessments and three dual-task assessments (Radomski et al., 2013). After initial pilot testing of the AMMP battery, data analysis indicated variable IRR (intraclass correlation coefficients [ICCs] of .45, .37, and .79 for task performance) for the three multitask assessments of executive functioning. Scoring was complicated by errors resulting from simultaneous observation and scoring requirements and by a lack of clearly defined scoring criteria outlining acceptable tolerances for partially accurate task performance. For example, when participants were told to obtain an address, rater disagreements occurred if part of the address was incorrect (e.g., transposed digits, spelling errors); some examiners gave full credit for task completion and others gave no credit. In addition to multiple scoring challenges, test developers indicated substantial test burden from three relatively similar multitask assessments and limited face validity of the tasks as reported by participants. In an effort to improve IRR, face validity, and clinical feasibility, the CQDT was developed to replace the three previous iterations of multitask assessments.

The first step in the development of the CQDT was to reexamine the literature pertaining to current multitask assessments. The team also shared the initial concept, materials, and instructions of the CQDT with a panel of experienced servicemembers who provided recommendations to improve face validity of the task with the target population. On the basis of the definition of multitasking (Burgess, 2000) and feedback from subject matter experts, the team created a list of parameters to be tested.

Once the initial task was developed, test developers practiced administering the task on servicemembers and civilians to observe variations in performance and variations in the interpretation of performance by multiple evaluators. After practice administrations, test developers clarified task instructions and revised the approach

to scoring by creating operational definitions that clarified situations in which no credit, partial credit, or full credit should be given. These operational definitions were included on the score sheet. For example, a participant who reported the incorrect number of barracks rooms would receive partial credit for task performance in that domain as determined by the operational definition for that task. This scoring approach reduced scoring complexity and allowed raters to assign a score quickly upon observation of task completion.

The score sheet was also improved to reduce scoring errors resulting from simultaneous observation and scoring requirements. Many aspects of the CQDT required scoring in real time (i.e., radio communications with various personnel on the correct radio frequency) to determine whether participants completed tasks independently and accurately or required cueing. Raters who were distracted or who failed to score performance on these tasks immediately made scoring errors. To address this issue, task assignments were listed chronologically on the score sheet, and tasks requiring immediate scoring were emphasized with bold font. This design helped cue the evaluators to ensure observation of performance at appropriate times. Last, the score sheet included correct responses for objective performance components (e.g., correct number of vacant barracks rooms to be reported, manufacturer's address), allowing the rater to quickly identify performance accuracy and assign the appropriate score. These additions were implemented to maximize scoring efficiency.

After all modifications were made to the CQDT, test developers piloted the revised multitask assessment in a healthy population to assess IRR. Given the anticipated variability in task performance between healthy servicemembers and those with concussion, evaluation of IRR in healthy servicemembers allowed for subsequent scoring and procedural refinements to be made before evaluating IRR in servicemembers with concussion.

### *Intrarater Reliability Testing*

Preliminary IRR was assessed between 3 (2 trained and 1 novice) raters when measuring individual participant performance

on the CQDT. The two trained raters were involved in test development, and the novice rater was a physical therapist with no prior experience with the CQDT. This design helped determine whether inexperienced providers could easily and accurately score the assessment. Before evaluating participants, the novice rater received a brief orientation (<30 min) to the score sheet, performance metrics, and operational definitions of task performance, rules, and rule breaks. IRR was established for all raters.

### *Participants*

Participants were recruited by convenience sampling from the U.S. Army Research Institute of Environmental Medicine in Natick, Massachusetts. All healthy active-duty servicemembers (active duty, guard, or reserve component) ages 18–42 yr were eligible to participate. Participants were excluded if they reported a history of traumatic brain injury (TBI) or concussion in the previous year, any documented active-duty restrictions (currently on a military profile), any physical or behavioral health condition preventing sustained activity for up to 30 min, history of psychiatric disorder, and uncorrected hearing deficits. All participants gave written informed consent before participation, and the institutional review board at the U.S. Army Research Institute of Environmental Medicine approved the study.

### *Data Collection*

The following components were measured via observation:

- *Task completion* was defined as the extent to which participants independently and accurately completed each assignment. Each assignment was scored 0 (*not complete*), 1 (*partially complete or required cueing to complete*), or 2 (*completed to defined standard independently without cueing*). The test included 17 assignments (some assignments required more than one task), with up to 2 points possible for each, for a total of 34 possible points for task completion.
- *Total rule breaks* for the four rules were operationally defined on the score sheet. Each rule that was broken was recorded.

- *Frequency of rule breaks* was recorded for each rule; it was possible to break the same rule multiple times. No limit was placed on the frequency of rule breaks.
- *Performance time* was defined as the total time to complete the task.
- *Transits* were defined as movements between work areas. Leaving one work area and entering another was considered one transit.

### *Data Analysis*

The ICC was used to quantify preliminary IRR. The Krippendorff (Hayes & Krippendorff, 2007)  $\alpha$  macro was run under SPSS Version 18.0 (IBM Corporation, Armonk, NY) to generate the ICCs. Twelve cases provided 95% confidence to measure our objective for an ICC of .90 against a minimum ICC of .70 (Bonett, 2002). For metrics that achieved an ICC of .90, the mean, standard deviation, and range are reported on the basis of the median of the three scores for each participant.

### *Results*

A total of 12 servicemembers (7 men and 5 women) participated in this study. The mean time to perform the CQDT was 19.6 min; 7 of 12 participants completed the task in <20 min and 11 of 12 in <23 min. The maximum test duration was 31.9 min. The average number of transits was 10.5. Table 1 provides the IRR results. Rule breaks and frequency of rule breaks were not reliable, with ICCs of .66 and .64, respectively. Task completion, transits, and total time were highly reliable, with ICCs of .94, .98, and .98, respectively.

### *Discussion*

Occupational therapists are charged with developing and implementing measurement strategies that characterize the extent to which impairments impede daily life performance (Baum, Perlmutter, & Dunn, 2005). Doing so is difficult when impairments such as executive dysfunction are potentially difficult to detect, as in servicemembers with concussion. Performance-based assessments that involve multitasking have demonstrated the potential to discriminate between

**Table 1. Preliminary Interrater Reliability Results for the Charge of Quarters Duty Task (N = 12)**

| Item                     | Reliability (ICC) | 95% CI     | Mean (SD)  | Range     |
|--------------------------|-------------------|------------|------------|-----------|
| Task completion          | .94               | [.86, .99] | 27.6 (5.6) | 13–33     |
| Rule breaks              | .66 <sup>a</sup>  | [.39, .88] |            |           |
| Frequency of rule breaks | .64 <sup>b</sup>  | [.32, .90] |            |           |
| Transits                 | .98               | [.96, .99] | 10.5 (4.0) | 5–18      |
| Total time (min)         | .98               | [.96, .99] | 19.6 (4.8) | 13.2–31.9 |

Note. CI = confidence interval; ICC = intraclass correlation coefficient; SD = standard deviation. The mean, standard deviation, and range are reported only for metrics that achieved an ICC of .90.

<sup>a</sup>Four of 12 triplets did not agree. <sup>b</sup>Six of 12 triplets did not agree.

healthy control participants and people with executive dysfunction (Alderman et al., 2003; Baum et al., 2008; Morrison et al., 2013; Wolf, Morrison, & Matheson, 2008) and may be an alternative to traditional measures of cognitive domains, which often fail to detect existing deficiencies in complex task performance (Tranel, Hathaway-Nepple, & Anderson, 2007). Although such tests do not appear to be subject to the ceiling effects of more structured measures of performance (Hall et al., 1996; Scott et al., 2011), they are typically complex to administer and score (Morrison et al., 2013). More multitasking tests that are specific to various clinical populations and life situations are needed. IRR specific to servicemembers with concussion and discriminant validity remain untested for the CQDT, but the preliminary evaluation of IRR in healthy participants suggests progress in the development of a multitask assessment of executive functioning for servicemembers with concussion.

The current evaluation of preliminary IRR highlights easily scored metrics for multitasking assessment and those requiring further refinement by the research team. IRR for task completion improved from previous versions of multitasking assessments because the score sheet was redesigned to include operational definitions and list performance tasks chronologically. These elements helped clarify scoring criteria and reduce rater disagreements regarding task performance.

Unfortunately, behavioral aspects of rule breaks and frequency of rule breaks were not as well specified, accounting for continued but solvable problems with IRR. Rater disagreements in how to score vocalizations directed at the examiners (e.g., asking the examiner questions) and the number of PVC parts brought to the assembly area largely explained the unacceptable ICCs for rule

breaks and frequency of rule breaks. Operational definitions were not clear enough to account for the unpredictable nature of human performance in these areas. Additionally, the restricted range resulting from only four rules may have had a negative impact on the ICC values. With a restricted range, one missed observation in rule breaks can affect the ICC value to a greater degree than with a greater number of rules. In preparation for future data collection, operational definitions have been revised and piloted to improve IRR for rule breaks.

## Limitations and Future Directions

The CQDT is in relative infancy in terms of test development. Thus far, clinical feasibility and IRR for the CQDT have been evaluated in only a small number of healthy participants. Results of future data collection will determine IRR and clinical feasibility of the CQDT in a clinical population and, most important, will ascertain whether it discriminates between healthy control participants and servicemembers with concussion. If so, further research will need to be conducted to determine whether the CQDT predicts successful return to duty. Finally, the team is exploring the development of a civilian version of the CQDT that could be used as a stand-alone assessment of executive dysfunction.

## Implications for Occupational Therapy Practice and Research

The results of this study have the following implications for occupational therapy practice and research:

- Performance-based assessments of multitasking may enable occupational therapy practitioners to identify executive function deficits after concussion.

- Because of the complexity of scoring a multitask assessment, operational definitions for scoring are best developed on the basis of observed variations in task performance and differences in interpretation of that performance by multiple evaluators.
- The lessons learned in the development of the CQDT may benefit occupational therapy practitioners interested in developing performance-based assessments of executive dysfunction tailored to populations and practice settings other than the military.

## Conclusion

There remains a need for reliable, valid, and clinically feasible assessments that can be used to identify executive dysfunction. Performance-based assessments that incorporate multitask methods and accurately simulate job demands may prove useful for occupational therapy practitioners in determining return-to-activity timelines in various populations. ▲

## Acknowledgments

This ongoing work was funded by the U.S. Army Medical Research and Materiel Command. We thank the soldiers who provided valuable feedback to improve the face validity of the Charge of Quarters Duty Task and the soldiers who participated in this study. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Army or the U.S. Department of Defense.

## References

- Alderman, N., Burgess, P. W., Knight, C., & Henman, C. (2003). Ecological validity of a simplified version of the Multiple Errands Shopping Test. *Journal of the International Neuropsychological Society*, 9, 31–44. <http://dx.doi.org/10.1017/S1355617703910046>
- Baum, C. M., Connor, L. T., Morrison, T., Hahn, M., Dromerick, A. W., & Edwards, D. F. (2008). Reliability, validity, and clinical utility of the Executive Function Performance Test: A measure of executive function in a sample of people with stroke. *American Journal of Occupational Therapy*,



- 62, 446–455. <http://dx.doi.org/10.5014/ajot.62.4.446>
- Baum, C. M., Perlmutter, M., & Dunn, W. (2005). Establishing the integrity of measurement data. In M. Law, C. Baum, & W. Dawn (Eds.), *Measuring occupational performance* (pp. 49–64). Thorofare, NJ: Slack.
- Bonett, D. G. (2002). Sample size requirements for estimating intraclass correlations with desired precision. *Statistics in Medicine*, 21, 1331–1335. <http://dx.doi.org/10.1002/sim.1108>
- Brenner, L. A., Terrio, H., Homaifar, B. Y., Gutierrez, P. M., Staves, P. J., Harwood, J. E., . . . Warden, D. (2010). Neuropsychological test performance in soldiers with blast-related mild TBI. *Neuropsychology*, 24, 160–167. <http://dx.doi.org/10.1037/a0017966>
- Burgess, P. W. (2000). Strategy application disorder: The role of the frontal lobes in human multitasking. *Psychological Research*, 63, 279–288. <http://dx.doi.org/10.1007/s004269900006>
- Carroll, L. J., Cassidy, J. D., Peloso, P. M., Borg, J., von Holst, H., Holm, L., . . . Pépin, M.; WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. (2004). Prognosis for mild traumatic brain injury: Results of the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. *Journal of Rehabilitation Medicine*, 43(Suppl.), 84–105. <http://dx.doi.org/10.1080/16501960410023859>
- Cuberos-Urbano, G., Caracuel, A., Vilar-López, R., Valls-Serrano, C., Bateman, A., & Verdejo-García, A. (2013). Ecological validity of the Multiple Errands Test using predictive models of dysexecutive problems in everyday life. *Journal of Clinical and Experimental Neuropsychology*, 35, 329–336. <http://dx.doi.org/10.1080/13803395.2013.776011>
- Dawson, D. R., Anderson, N. D., Burgess, P., Cooper, E., Krpan, K. M., & Stuss, D. T. (2009). Further development of the Multiple Errands Test: Standardized scoring, reliability, and ecological validity for the Baycrest version. *Archives of Physical Medicine and Rehabilitation*, 90(Suppl.), S41–S51. <http://dx.doi.org/10.1016/j.apmr.2009.07.012>
- Frisch, S., Förstl, S., Legler, A., Schöpe, S., & Goebel, H. (2012). The interleaving of actions in everyday life multitasking demands. *Journal of Neuropsychology*, 6, 257–269. <http://dx.doi.org/10.1111/j.1748-6653.2012.02026.x>
- Hall, K. M., Mann, N., High, W. M., Wright, J., Kreutzer, J. S., & Wood, D. (1996). Functional measures after traumatic brain injury: Ceiling effects of FIM, FIM+FAM DRS and CIQ. *Journal of Head Trauma Rehabilitation*, 11, 27–39. <http://dx.doi.org/10.1097/00001199-199610000-00004>
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1, 77–89. <http://dx.doi.org/10.1080/19312450709336664>
- Helmick, K., Baugh, L., Lattimore, T., & Goldman, S. (2012). Traumatic brain injury: Next steps, research needed, and priority focus areas. *Military Medicine*, 177(Suppl.), 86–92. <http://dx.doi.org/10.7205/MILMED-D-12-00174>
- McCrea, M., Iverson, G. L., McAllister, T. W., Hammeke, T. A., Powell, M. R., Barr, W. B., & Kelly, J. P. (2009). An integrated review of recovery after mild traumatic brain injury (MTBI): Implications for clinical management. *Clinical Neuropsychologist*, 23, 1368–1390. <http://dx.doi.org/10.1080/13854040903074652>
- Morrison, M. T., Giles, G. M., Ryan, J. D., Baum, C. M., Dromerick, A. W., Polatajko, H. J., & Edwards, D. F. (2013). Multiple Errands Test–Revised (MET–R): A performance-based measure of executive function in people with mild cerebrovascular accident. *American Journal of Occupational Therapy*, 67, 460–468. <http://dx.doi.org/10.5014/ajot.2013.007880>
- Radomski, M. V., & Morrison, M. T. (2014). Assessing abilities and capacities: Cognition. In *Occupational therapy for physical dysfunction* (7th ed., pp. 121–143). Baltimore: Lippincott Williams & Wilkins.
- Radomski, M. V., Weightman, M. M., Davidson, L. F., Finkelstein, M., Goldman, S., McCulloch, K., . . . Stern, E. B. (2013). Development of a measure to inform return-to-duty decision making after mild traumatic brain injury. *Military Medicine*, 178, 246–253. <http://dx.doi.org/10.7205/MILMED-D-12-00144>
- Scott, J. C., Woods, S. P., Vigil, O., Heaton, R. K., Schweinsburg, B. C., Ellis, R. J., . . . Marcotte, T. D.; San Diego HIV Neurobehavioral Research Center Group. (2011). A neuropsychological investigation of multitasking in HIV infection: Implications for everyday functioning. *Neuropsychology*, 25, 511–519. <http://dx.doi.org/10.1037/a0022491>
- Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, 114, 727–741. <http://dx.doi.org/10.1093/brain/114.2.727>
- Tranel, D., Hathaway-Nepple, J., & Anderson, S. W. (2007). Impaired behavior on real-world tasks following damage to the ventromedial prefrontal cortex. *Journal of Clinical and Experimental Neuropsychology*, 29, 319–332. <http://dx.doi.org/10.1080/13803390600701376>
- Vagnozzi, R., Signoretti, S., Tavazzi, B., Floris, R., Ludovici, A., Marziali, S., . . . Lazzarino, G. (2008). Temporal window of metabolic brain vulnerability to concussion: A pilot <sup>1</sup>H-magnetic resonance spectroscopic study in concussed athletes—Part III. *Neurosurgery*, 62, 1286–1295, discussion 1295–1296. <http://dx.doi.org/10.1227/01.neu.0000333300.34189.74>
- Wolf, T. J., Morrison, T., & Matheson, L. (2008). Initial development of a work-related assessment of dysexecutive syndrome: The Complex Task Performance Assessment. *Work*, 31, 221–228.